LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-JRNL-562205

Advection of Surface-Derived Organic Carbon Fuels Microbial Reduction in Bangladesh Groundwater

B. J. Mailloux, E. Trembath-Reichert, J. Cheung, M. Watson, M. Stute, G. A. Freyer, A. Ferguson, K. M. Ahmed, J. Alam, B. A. Buchholz, J. Thomas, A. Layton, Y. Zheng, B. C. Bostick, A. van Geen

June 22, 2012

Proceedings of the National Academy of Sciences

Disclaimer

Title: Advection of Surface-Derived Organic Carbon Fuels Microbial Reduction in Bangladesh Groundwater

Authors: Brian J. Mailloux[1*], Elizabeth Trembath-Reichert[1], Jennifer Cheung[1], Marlena Watson[1], Martin Stute[1], Greg A. Freyer[2], Andrew Ferguson[3], Kazi Matin Ahmed[4], Md, J. Alam[4], Bruce A. Buchholz[5], James Thomas[6], Alice Layton[7], Yan Zheng[8], Benjamin C. Bostick[8], Alexander van Geen[8]

Affiliations:
[1]Environmental Science Department, Barnard College, NY, NY 10027

[2]Environmental Health Science, Columbia University, NY, NY 10032

[3]Department of Civil Engineering and Engineering Mechanics, Columbia University, NY, NY 10027

[4]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964

[4]Department of Geology, University of Dhaka, Dhaka 1000, Bangladesh
[5]Center for Accelerator Mass Spectrometry, Lawrence Livermore National Laboratory, Livermore, CA 94551-9900

[6]Hydrologic Sciences, Desert Research Institute, Reno, NV, 89512

[7]Department of Microbiology and Center for Environmental Biotechnology, U Tennessee Knoxville, Knoxville, TN, 37922
[8]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964

*Correspondence to: bmaillou@barnard.edu

Abstract: Chronic exposure to arsenic (As) by drinking shallow groundwater pumped from tubewells causes widespread disease in Bangladesh and neighboring countries from South to East Asia (*1*). The release of As naturally present in the sediment to groundwater has been linked to reductive dissolution of iron oxides coupled to the microbial respiration of reactive organic carbon (OC) (*2*). The source of reactive OC driving this microbial reduction is still debated despite its importance in regulating aquifer redox status, and thereby aqueous As levels (*3*). We show here that analyzing the radiocarbon (^{14}C) signature of microbial DNA isolated from aquifer microbes is a novel method uniquely suited to determine the relative importance of surface and aquifer-derived OC sources for microbial reduction. DNA samples collected from the shallow, high-As aquifer and one sample from the underlying, low-As aquifer were consistently younger than the sediment, by as much as several thousand years. This difference implies that microbes preferentially utilize younger, surface-derived OC pools. The vertical profile shows that downward transport of dissolved OC is occurring on anthropogenic timescales, but bomb ^{14}C-labelled dissolved OC has not yet accumulated in DNA and is not fueling reduction. These results confirm that As release is a natural process that pre-dates human perturbations to groundwater, even if such perturbations could affect groundwater As levels in the future.

Main Text: Aquifer redox status is a major factor in determining groundwater quality and potability. The most egregious example is the dire health impact of elevated levels of As in groundwater drawn with inexpensive hand-pumped tubewells by more than 100 million villagers across South, Southeast, and East Asia (*1, 4*). Aquifer redox status is largely controlled by microbial respiration of organic carbon (OC). Heterotrophic reduction of iron (Fe) oxides containing As coupled with the oxidation of OC is a leading cause of the accumulation of As to groundwater (*1*), but there is no consensus on the source of OC. Such information is critical to understanding aquifer dynamics and providing safe drinking water options.

Human perturbations have long been suggested to exacerbate the problem of arsenic in groundwater. Given the connectivity between contaminated surface waters and aquifers, widespread pumping could redistribute As from the surface to depth (*5, 6*) or between aquifers (*7*). This pumping could potentially also deliver reactive carbon from ponds and latrines to depth, causing microbial reduction and As release (*5, 6*). Fluorescent spectra of dissolved organic carbon (DOC), correlations of sedimentary organic carbon (SOC) and arsenic levels, and stratigraphy suggest, on the other hand, that the reactive OC is derived from the sediment and has a geologic origin (*8-10*). Surprisingly, given the importance of the issue for the health of such a large population, no attempt has been made to unequivocally determine the source of reactive OC in aquifers of this region or worldwide. Here, we directly identify the OC source being utilized by microorganisms in a high-As aquifer.

RADIOCARBON ANALYSIS of DNA

We have developed a method to analyze the ambient radiocarbon (^{14}C) signature of microbial DNA collected from groundwater. DNA is mainly present in viable cells and synthesized during growth utilizing carbon from the surrounding environment. Therefore, the radiocarbon signature of DNA becomes a direct measure of the carbon

utilized during microbial respiration and growth and can be used to determine the source of OC fueling microbial reduction and replication. Radiocarbon analysis has been performed on DNA extracted from human cells (*11*) and, attempted on surface water samples (*12, 13*). The method has never been applied to study bacteria collected from groundwater. The difficulty with radiocarbon analysis of DNA from environmental samples lies in collecting and purifying enough DNA (>100μg DNA) without introducing exogenous carbon contamination. We have overcome these limitations by concentrating bacteria from over 3000 L (equivalent to a 1.8m radius around a 1m wellscreen) of groundwater by filtering onto a 0.2 μm nylon filter. DNA is physically and chemically extracted from the filter, purified, and radiocarbon dated (Supplemental information). In control samples, the blank filter contained no DNA and there was close agreement between the radiocarbon age of the media and the DNA extracted from cells grown on known carbon sources (Table 1, Supplemental information). The results indicate that radiocarbon measurements for DNA extracted from Bangladesh groundwater can be compared with radiocarbon data for sediment, groundwater dissolved organic carbon (DOC), and dissolved inorganic carbon (DIC) to determine sources of OC used for microbial assimilation.

RADIOCARBON in DNA from BANGLADESH SAMPLES

Here, we report a depth profile of DNA radiocarbon measurements at a well-studied site in Bangladesh where groundwater is pumped on a large scale during the dry season for irrigation (Figure 1). The study site (Site F) has been the focus of extensive investigations and is located in Lashkardi village of Araihazar upazila, approximately 20 km east of the capital Dhaka (*14-16*). The water table at the site rises by over 3 m during the summer monsoon. The surface soil and sediment surrounding the village is particularly sandy (*17*) and the area contains a large number of high-capacity irrigation wells (Figure 1). Because of these characteristics the groundwater recharge rate of 0.5 m/y (corresponding to a vertical velocity of 2 m/y near the water table, porosity = 0.25) inferred from ^{3}H-^{3}He groundwater dating is particularly high (*15*). The imprint of human activities on biogeochemistry, if significant, is likely to be detected at this site.

The vertical profile in Lashkardi village represents the evolution of groundwater geochemistry from the surface, although it does not directly trace a groundwater flow path (*18*). The five shallow wells ranging between 6 and 25 m in depth below ground tap a Holocene (<10,000 year-old) aquifer that has developed the characteristic vertical As gradient with As concentrations increasing from 0.007 to 2.7 μM (<1-203 μg/l) (*14, 16*). Within the same depth interval, concentrations of Fe, methane (CH_4) and DIC increase while redox potential (Eh) and sulfate concentrations decrease, consistent with microbial respiration of OC coupled to reductive dissolution of Fe oxides (Figure 2). Low-levels (<0.25%) to non-detectable sedimentary carbonates indicates that calcite dissolution and precipitation are insignificant. The DIC thus represents the sum of atmospheric carbonate in recharge and respired organic matter and bounds the maximum water age. There is a thin clay layer separating the two wells at 20 and 25 m depth within the Holocene aquifer. Above this thin clay layer, the aquifer contains ^{3}H and has been recharged since atmospheric testing of nuclear bombs began in the 1950's whereas deeper groundwater has no ^{3}H and contains older (>50y) groundwater. The deepest monitoring well extends beyond 15m of clay and reaches a Pleistocene (>10,000 y) aquifer at 57 m containing

orange oxidized sands and groundwater with <5 μg/L As that is likely much older than the shallower water (Figure 1, Supplemental Figure 1). The radiocarbon signature of the microbial DNA was measured in four samples, three from the Holocene and one from the Pleistocene.

Starting from the Pleistocene aquifer, the radiocarbon age of microbial DNA at 57m is 4,680±60 ^{14}C y, which is equivalent to Δ^{14}C -446±4‰ or 0.56 Fraction modern (Fm) in other notations. This DNA is considerably younger than the Pleistocene sediment. Therefore the carbon utilized by the resident microbes does not originate solely from the sediment and a significant portion has to be supplied from younger, exogenous sources through advection. Advection also explains why DOC is considerably younger than the sediment. In fact, the DNA age is similar to that of DOC and DIC, which were dated at 5,560±42 ^{14}C y (-502±3‰) and 6240±30 ^{14}C y (-543‰), respectively.

The age of the OC in the sediment corresponding to the next deepest Holocene interval in the profile (25m) is about 9,200 ^{14}C y and contrasts with that of the groundwater which is >50 y old (based on lack of ^{3}H) and <540±25 y (based on DIC ^{14}C age, Figure 3, SI Table). In this depth interval, the DNA (1,185±40 ^{14}C y or Δ^{14}C -143±4‰) also is much younger than the sediment and also much closer to the age of the DOC (865±35 ^{14}C y) (Figures 2 and 3). The deeper depth intervals (25 and 57m) of the profile provide new constraints on the source of reactive carbon driving reduction. The DNA ages suggests that advected OC contributes to the reactive OC pool. This surprising result indicates that carbon is effectively advected to depth, even when separated from the surface clay.

Metagenomic analysis of DNA from the 25m well allows us to more completely characterize the microbial populations and interpret the radiocarbon data. The data indicate that the microbial community is dominated by species and subsystems associated with redox reactions and degradation of complex carbon sources (Supporting Information). However, methanogenesis could potentially lead to the incorporation of DIC into DNA, making the DNA ages more similar to the younger DIC. Methane concentrations peak at 25m, but autorophic methanogens belonging to Archaea represented only 3% of the total sequences within this zone. In contrast Proteobacteria which are predominantly heterotrophic represented >50% of the total sequences (supplemental information). Thus, it is unlikely that methanogenesis strongly affects measured DNA ages. Heterotrophic metal-reducing organisms thus appear to dominate within the aquifer. Paired with the younger and similar ^{14}C ages of DNA and DOC relative to sedimentary OC, we conclude that advected OC is sustaining microbial reduction in this system.

Groundwater from depths <19 m containing bomb-produced ^{3}H is most likely to reflect human perturbation. The presence of a strong bomb-^{14}C spike (Δ^{14}C 370±5‰; Fig. 2) for DOC at 11m confirms that young DOC has been transported to this depth within approximately the last 30 to 50 years. In contrast, there is no indication of a significant bomb-carbon input in the DOC at 19m (Δ^{14}C -205±4‰; Fig. 2). These observations combined indicate that the advection of DOC is delayed relative to that of groundwater. We conservatively estimate a retardation factor for DOC transport of 5.5 by dividing the ratio of the vertical groundwater velocity at this site (Stute et al. 2007) to the fastest potential vertical penetration of bomb-labeled DOC (11 m in 30 years). Not only

is DOC retarded relative to advection, its utilization by resident microbes appears to be quite limited. The low ^{14}C content of the DNA extracted at 11m (410±90 ^{14}C y or $\Delta^{14}C$ -56±10‰) indicates that little bomb-spiked DOC has been metabolized by the resident microbes.

The radiocarbon age of DNA at 19m (940±50 ^{14}C y or $\Delta^{14}C$ -117±5‰) is younger than the sediment and intermediate and collinear between DNA ages at 11 and 25m (Figure 3). The combination of all three ages extrapolates to approximately zero age at the water table surface and therefore indicates an effective rate of reactive OC transport from the surface of 0.018 m/y. This is over 100-fold slower than the vertical groundwater velocity, and roughly 20-fold slower than the estimated DOC velocity. Thus, advected surface-derived OC is not quickly consumed and can persist for decades without being metabolized in the aquifer.

IMPLICATIONS

The generation, delivery, and transformation of surficial organic carbon (OC) regulates aquifer redox status and the utilization of this OC plays a central role in regulating arsenic levels at this site. Although advected OC is used in the aquifer, the new data reported here challenge the notions that young OC drawn down by irrigation pumping over the last 50 years has affected water quality in a rapidly recharged aquifer (*5, 6*). In addition, the data place strong limits on the role that very old OC contained in petroleum (*19*), isolated peat strata (*8*) or ambient sedimentary carbon (*9*) play in driving microbial activity in reducing aquifers. In all sampled intervals of Holocene and Pleistocene age, the DNA is considerably younger than the sediment. Advection of OC from a younger interval, most likely very shallow soils based on the extrapolated depth trend of DNA age, plays a central role. These data cannot explain why DOC concentrations increase with depth but show that the reactive OC is not from recent anthropogenic recharge at all depths. The shallow OC that recharges the aquifer, remains labile, is metabolized slowly by the microbial communities at depth, and is retarded, presumably by adsorption, relative to groundwater flow. Given the retardation and utilization, OC from the last hundred years reaches only a few meters into the sediments and is not associated with elevated As levels.

These results do not imply that recharge enhancement in areas of intense irrigation pumping have not affected the local hydrology and As transport through readsorption, flushing, or other mechanisms (*20, 21*). Our observations also do not rule out the possibility that enhanced recharge of reactive DOC could lead to more reductive dissolution of Fe oxides in the future (*3*), particularly if pumping rates accelerate (*22*). These results suggest that arsenic release to date is primarily a natural process fueled by the slow advection of surface derived OC and that the advection of OC including anthropogenic sources might affect As but this has not yet occurred. Thus, low-As supplies that are isolated from organic rich recharge should be the preferred long-term drinking water source.

References and Notes:

1. P. Ravenscroft, H. Brammer, K. Richards, *Arsenic Pollution*. (Wiley-Blackwell, 2009).

2. R. Nickson *et al.*, *Nature* **395**, 338. (1998).
3. S. Fendorf, H. A. Michael, A. van Geen, *Science* **328**, 1123 (2010).
4. M. Argos *et al.*, *The Lancet* **376**, 252 (2010).
5. C. F. Harvey *et al.*, *Science* **298**, 1602 (Nov 22, 2002).
6. R. B. Neumann *et al.*, *Nature Geosci* **3**, 46 (2010).
7. K. A. Radloff *et al.*, *Nature Geosci* **4**, 793 (2011).
8. J. M. McArthur *et al.*, *Applied Geochemistry* **19**, 1255 (2004).
9. A. Meharg *et al.*, *Environmental Science & Technology* **40**, 4928 (2006).
10. N. Mladenov *et al.*, *Environmental Science & Technology* **44**, 123 (2009).
11. K. L. Spalding, R. D. Bhardwaj, B. A. Buchholz, H. Druid, J. Frisen, *Cell* **122**, 133 (2005).
12. J. Cherrier, J. E. Bauer, E. R. M. Druffel, R. B. Coffin, J. P. Chanton, *Limnology and Oceanography* **44**, 730 (1999).
13. R. L. Hansman *et al.*, *Proceedings of the National Academy of Sciences* **106**, 6513 (2009).
14. A. Horneman *et al.*, *Geochimica et Cosmochimica Acta* **68**, 3459 (2004).
15. M. Stute *et al.*, *Water Resour. Res.* **43**, W09417 (2007).
16. R. K. Dhar *et al.*, *J. Contam. Hydrol.* **99**, 97 (2008).
17. Z. Aziz *et al.*, *Water Resour. Res.* **44**, 15 (Jul, 2008).
18. J. C. Vogel, in *Isotopes in Hydrology*. (IAEA, Vienna, 1967), vol. IAEA-SM-83/24.
19. H. A. L. Rowland, D. A. Polya, J. R. Lloyd, R. D. Pancost, *Organic Geochemistry* **37**, 1101 (2006).
20. A. van Geen *et al.*, *Environmental Science & Technology* **42**, 2283 (2008).
21. J. M. McArthur *et al.*, *Water Research* **44**, 4171 (2010).
22. H. A. Michael, C. I. Voss, *Proceedings of the National Academy of Sciences* **105**, 8531 (2008).
23. A. van Geen *et al.*, *Water Resour. Res.* **39**, 1140 (2003).
24. S. L. Goodbred Jr, S. A. Kuehl, *Sedimentary Geology* **133**, 227 (2000).
25. Y. Zheng *et al.*, *Applied Geochemistry* **19**, 201 (2004/2, 2004).
26. S. L. Goodbred Jr, S. A. Kuehl, *Sedimentary Geology* **121**, 239 (1998).
27. G. S. Burr *et al.*, *Radiocarbon* **43**, 177 (2001).
28. A. Horneman *et al.*, *J. Contam. Hydrol.* **97**, 27 (2008).
29. Y. Zheng *et al.*, *Geochimica et Cosmochimica Acta* **69**, 5203 (2005/11/15, 2005).
30. J. a. R. Sambrook, David, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, ed. Third, 2001), vol. 3.
31. J. S. Vogel, J. R. Southon, D. E. Nelson, *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms* **29**, 50 (Nov, 1987).
32. G. M. Santos, J. R. Southon, K. C. Druffel-Rodriguez, S. Griffin, M. Mazon, *Radiocarbon* **46**, 165 (2004).
33. T. A. Brown, J. R. Southon, *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **123**, 208 (1997).
34. M. Stuiver, H. A. Polach, *Radiocarbon* **19**, 355 (1977).

35. M. L. Sogin *et al.*, *Proceedings of the National Academy of Sciences* **103**, 12115 (August 8, 2006, 2006).

36. V. Acosta-Martínez, S. Dowd, Y. Sun, V. Allen, *Soil Biology and Biochemistry* **40**, 2762 (2008).

37. T. Z. DeSantis *et al.*, *Microb. Ecol.* **53**, 371 (Apr, 2007).

38. M. Hamady, J. J. Walker, J. K. Harris, N. J. Gold, R. Knight, *Nat Meth* **5**, 235 (2008).

39. S. M. Huse *et al.*, *PLoS Genetics* **4**, 1 (2008).

40. F. Meyer *et al.*, *BMC Bioinformatics* **9**, 386 (2008).

41. N. Abu Laban, D. Selesi, T. Rattei, P. Tischler, R. U. Meckenstock, *Environmental Microbiology* **12**, 2783 (2010).

42. R. T. Anderson, J. N. Rooney-Varga, C. V. Gaw, D. R. Lovley, *Environmental Science and Technology* **32**, 1222 (1998).

43. R. Overbeek *et al.*, *Nucleic Acids Research* **33**, 5691 (2005).

Acknowledgments: All data are in the SI. We are grateful for the help of J. Mey, R. Foster, G. O'Mullan, P. Knappett, G. F. Slater, L. A. Ziolkowski, A. Dochenetz, and T. C. Onstott. This workwas performed in part under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. This work was supported by the Superfund Research Program (NIEHS Grant ES010349) and The Earth Institute at CU.

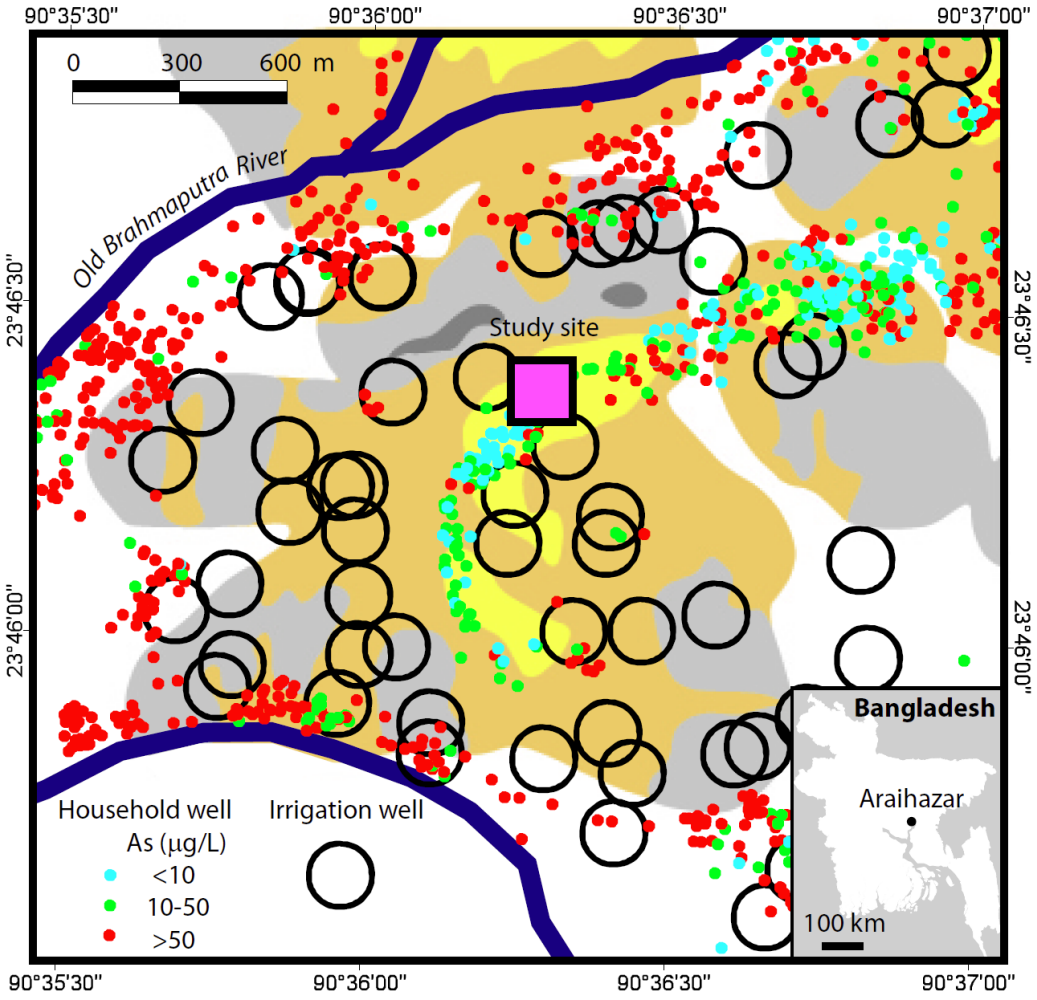

Fig. 1. Map of study location in Bangladesh (inset) and area surrounding study location at Site F. Small circles indicate As concentrations in private tubewells and delineate the villages (*23*). Larger circles represent high capacity irrigation wells (*14*). The background color is interpolated EM31 resistivity data (*17*) and the yellow shades represent more permeable surface sediments and gray shades represent fine-grained surface sediment.

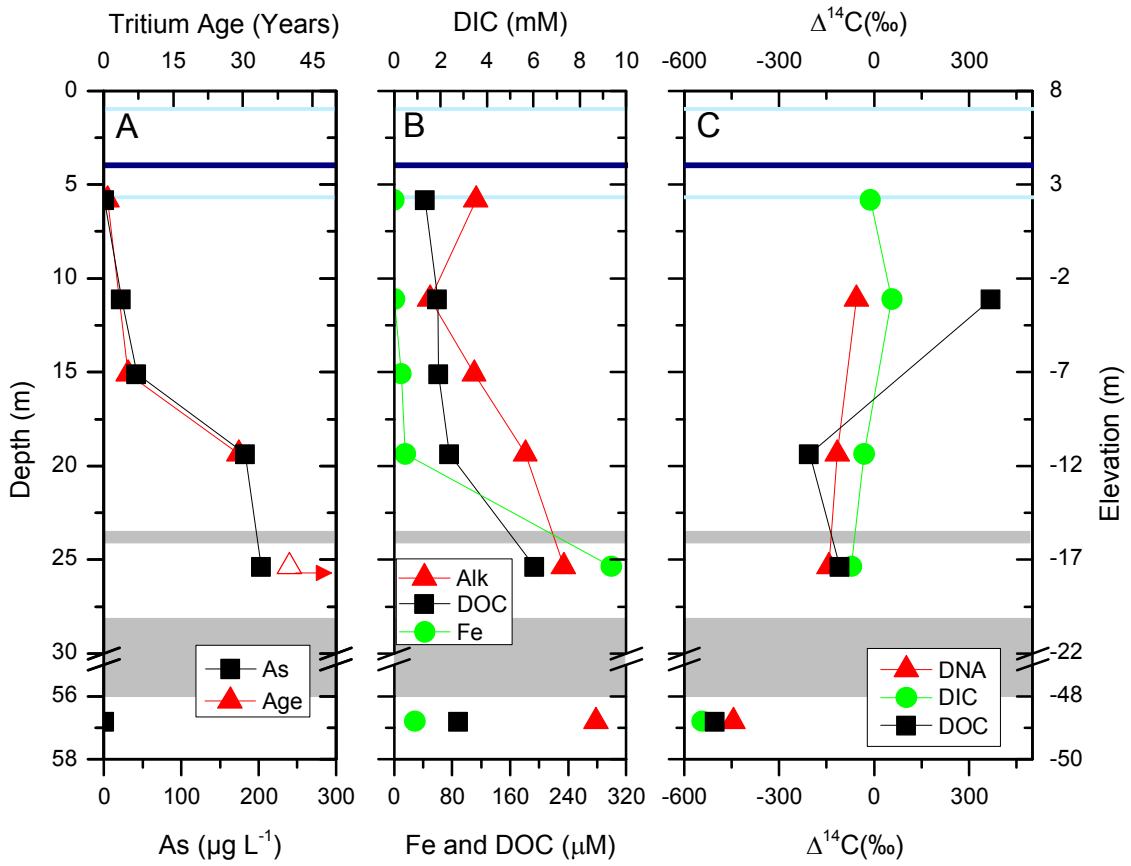

Fig. 2. Depth profiles of aqueous parameters from the Site F well nest in Lashkardi village, Bangladesh. There is a break in the depth from 30-55m which is a continuous clay/silt unit. Silt/clay units are represented by a gray background. The dark blue line at 4.0m on each panel is the average water table elevation and the lighter blue lines are the observed minimum and maximum water table elevations. A) Arsenic concentrations and Tritium ages of the water (*15*). The sample at 25m was greater than 40 years old and is represented by an open triangle with an arrow. B) Dissolved inorganic Carbon (DIC), total Fe and dissolved organic carbon concentrations. C) $\Delta^{14}C$ (‰) of DNA, DOC, and DIC. Data are also presented in years and Fm in supplemental information.

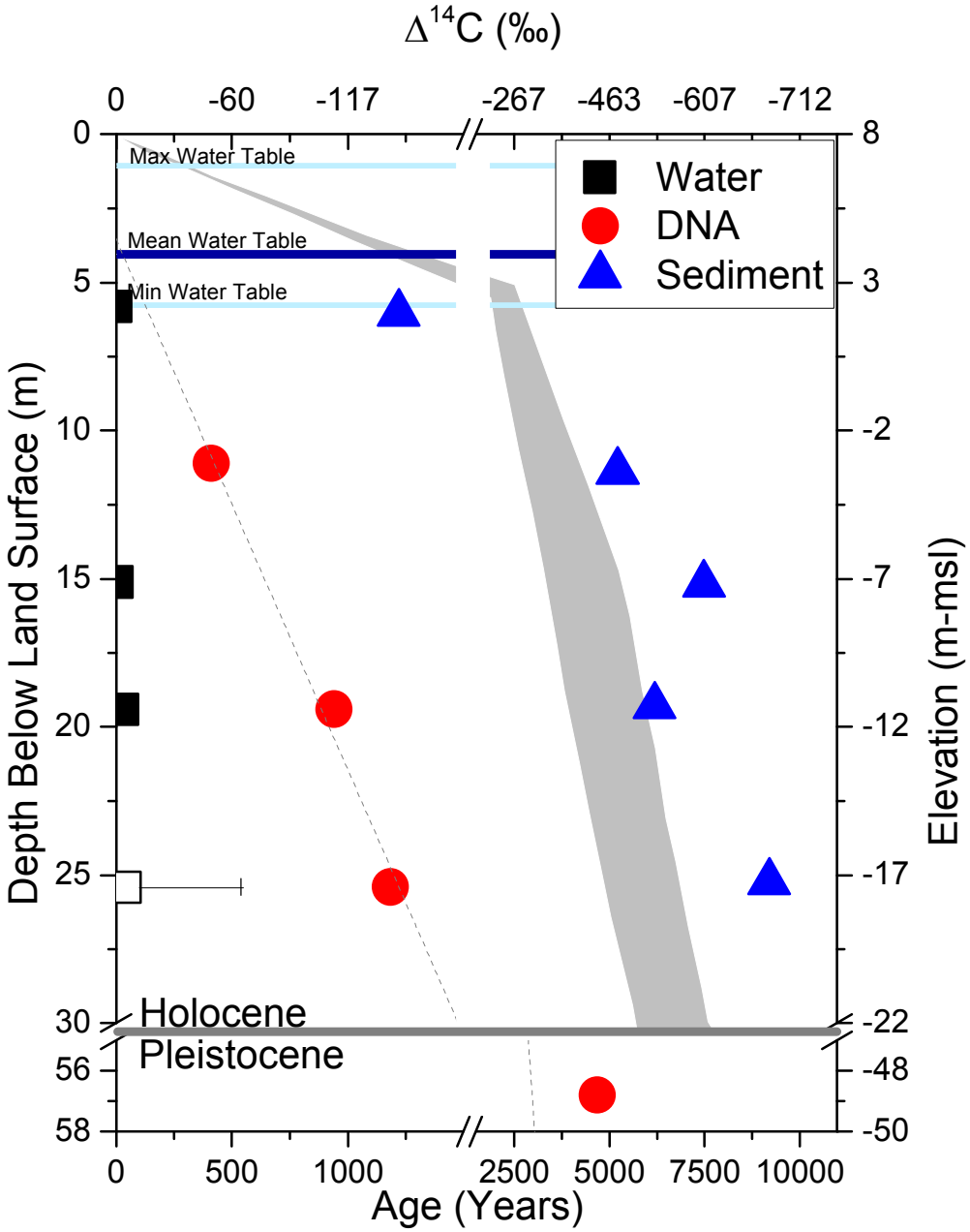

Fig 3. Vertical profiles comparing water ages, DNA ages, and sediment ages. Note that there is a break in the depths from 30-55m and the scale on the ages changes at 1500 years. The water ages are based on tritium-helium ($^3H/^3He$) groundwater dating. The sample at 25m contained no tritium and is older than 40 years but less than 540 years based on DIC. The sediment ages are bulk analyses of sediment cores. The gray zone represents predicted ages based on a depositional model (*24*). The dashed line is the best fit interpolation of the 3 shallowest DNA ages.

Table 1. Summary of radiocarbon data for DNA. nd indicates not detected and blank cells represent where data could not be measured. Errors are reported AMS errors.

	CAMS#	Volume Filtered (l)	A260/280	A260/230	$\Delta^{14}C$ ‰	Radiocarbon age (y)
Blank filter			nd	nd		
LB media	143162		-	-	54±4	Modern
LB-E. Coli	143159		1.88	2.33	6±6	Modern
Acetate	149428				-860±1	15715±45
Acetate DNA- Jenn	151193		1.82	2.13	-849±3	15170±170
F2-11.1m	141040	3,780	1.93	2.02	-56±10	410±90
F4-19.4m	143158	11,103	1.83	2.12	-117±5	940±50
F6-25.4,	147445	24,000	1.87	2.19	-143±4	1185±40
F5-56.8m	147444	17,464	1.86	2.24	-446±4	4680±60

Supplementary Materials:

Supplementary Materials:

Site Location and Characterization

Site F is located in Lashkardi village (23.774 °N, 90.606 °E) where nearly all the existing private tube wells contain less than 50 μg/l As (*14*). The shallow aquifer has an average grain size of 220±70 μm (Supplemental Table 1) and the maximum As concentration of 2.7 μM (203 μg/l) is reached at 25 m depth. The underlying clay layer separating the shallow aquifer from the deeper Pleistocene Dupi Tila aquifer extends from 32 to 47 m. Two organic rich horizons from the clay were sampled from drilling cuttings using the hand flapper method at 38 m (125 ft) (NOSAMS OS-91713, Fraction modern is 0.2931±0.0015, Age of 9860±40) and 39m (128 ft) (NOSAMS OS-91678, Fraction modern is 0.2898±0.0016, Age of 9950±45) and analyzed for radiocarbon content. These data suggests that below the silt/clay layer is Pleistocene sediment, consistent with other studies in the area (*25*). The four ages for the shallow Holocene aquifer range from 1220±75 y at 5.94 m to 9200±110 y at 25.15 m. These ages are consistent with infilling of the basin during Holocene deglaciation and sea-level rise and a subsequent decrease in deposition rate towards present day (*26*).

Supplemental Table 1. Sediment and radiocarbon data from site F.

Depth (m)	Depth (ft)	Mean Grain Size (mm)	Median Grain Size, D50 (mm)	Percent <63µm	Percent <2µm	LOI organic C (%)	$CaCO_3$ (%)	Accession #	Fraction Modern	Radiocarbon Age (y)	$\delta^{13}C$(‰)	$\Delta^{14}C$(‰)
5.6	19	0.294	0.369	11.0%	0.53%	NA	NA					
5.9	20	0.096	0.184	36.7%	3.9%	1.61	0.25	OS-82038	0.859±0.008	1220±75	ND	-147±8
9.0	30	0.196	0.298	17.2%	0.9%	NA	NA					
11.3	37	0.271	0.417	16.6%	1.1%	0.48	ND	OS-82037	0.522±0.003	5220±50	-27.8	-482±3
15.1	50	0.282	0.431	19.3%	1.1%	0.45	0.02	OS-82133	0.394±0.002	7480±45	-29.25	-609±1
19.2	63	0.232	0.328	14.9%	0.9%	0.38	0.01	OS-82036	0.463±0.003	6190±45	-27.85	-541±3
25.1	83	0.220	0.344	19.0%	1.5%	0.55	ND	OS-82217	0.318±0.004	9200±110	-24.66	-684±4
25.5	84	0.181	0.305	22.2%	1.6%	NA	NA					

Site F has been the focus of previous studies and aqueous species were collected and analyzed as part of these studies (*14-16*). Repeat analyses indicates that the groundwater parameters have been steady over time (*16*). DOC samples were collected from water after it passed through the DNA filters near the end of the sampling interval (Supplemental Table 2). DOC samples were analyzed at the University of Arizona using standard procedures (*27*). DIC samples were collected and analyzed according to standard protocols at NOSAMS. http://www.whoi.edu/nosams/page.do?pid=40135 . Samples were collected and analyzed during initial sampling (F1 and F5) (*15, 16*) and then during DNA sampling (F2, F4, and F6) (Supplemental Table 3). All radiocarbon data are presented with the reported AMS errors. Aqueous geochemistry data from the site that has been presented in previous studies and has been utilized in this study (Supplemental Table 4) (*14-16, 28*). Methods for aqueous samples are presented in Zhang et al., (*29*).

Supplemental Table 2. DOC radiocarbon Data.

Well	Well Depth (m)	$\Delta^{14}C$ (‰)	Radiocarbon Age (y)	Fraction Modern
F2	11.1	370±5	Modern	1.380±0.005
F4	19.4	-205±4	1759±38	0.803±0.004
F6	25.4	-108±4	865±35	0.898±0.004
F5	56.8	-502±3	5560±42	0.501±0.003

Supplemental Table 3. DIC radiocarbon Data.

Well	Well Depth (m)	$\delta^{13}C$(‰)	$\Delta^{14}C$ (‰)	Radiocarbon Age (y)	Fraction Modern	Accession Number
F1	5.8	-18.73	-12±4	45±30	0.994±.004	OS-41274
F2	11.1	-19.85	56±5	modern	1.063±.004	OS-82136
F4	19.4	-16.72	-31±4	195±25	0.976±.003	OS-82134
F6	25.4	-16.89	-72±6	540±25	0.935±.003	OS-82135
F5	56.8	-25.92	-543±2	6240±30	0.460±.002	OS-41275

Supplemental Table 4. Aqueous geochemistry data.

Well	Well Depth (m)	Well Elevation (m-msl)	CU sample # GB	Sampling Date and Time	T (˚C)	Conductivity (mS m^{-1})	pH	O2 (mg l^{-1})	Eh (mV)	Alkalinity (meq l^{-1})	DIC (mM)	H_2 (nM)	CH_4 (µM)	tritium age (y)
F1	5.8	2.08	605	1/7/03 11:11	26.3	33.8	6.30	nd	112	1.56	3.52	3.3	2.3	0.79
F2	11.1	-3.25	604	1/6/03 16:23	26.0	18.2	6.34	nd	-20	0.72	1.55	na	na	na
F3	15.1	-7.29	603	1/6/03 12:57	26.1	27.2	6.83	nd	-71	2.51	3.44	2.1	2.2	5.28
F4	19.4	-11.60	602A	1/6/03 13:20	26.1	46.0	6.92	nd	-96	4.34	5.65	2.5	0.6	29.10
F6	25.4	-17.46	601A	1/6/03 12:01	26.2	63.0	7.00	nd	-200	5.83	7.30	9.7	79.3	nd
F5	56.8	-49.0641	606	1/7/03 12:15	26.2	188.0	6.94	nd	-40	6.74	8.69	na	na	nd

Well	NH_4 (mg l^{-1})	Si (µM)	P (µM)	S (µM)	Cl (mM)	Na (mM)	Mg (mM)	K (mM)	Ca (mM)	Mn (µM)	Fe (µM)	As (µg l^{-1})	Al (µM)	DOC (mg l^{-1})
F1	6	573	0.2	65.2	0.64	0.46	0.51	0.07	0.56	0.14	0.15	0.54	0.33	0.52
F2	8	529	0.2	39.1	0.13	0.20	0.35	0.07	0.36	20.89	0.35	22.47	0.30	0.72
F3	18	4760	3.6	16.7	0.06	0.19	0.35	0.07	0.02	22.42	9.65	42.38	0.37	0.74
F4	39	6518	14.4	43.0	0.09	0.44	0.42	0.09	0.03	42.05	15.52	182.72	3.30	0.91
F6	147	7558	28.8	4.2	0.13	0.61	0.54	0.11	0.84	31.10	299.58	203.24	8.95	2.32
F5	33	7519	4.2	16.6	8.78	13.88	0.27	0.05	0.03	8.86	28.78	1.04	0.42	1.07

DNA Collection and Analysis

Filtering and Pumping

Water was collected and filtered using submersible plastic pumps (Groundwater Essentials, Sarasota, FL), 10" filter housing (Cole-Parmer) and 0.2 μm Serial Nylon filters (Cole-Parmer). Water was pumped through the filters for between 24 and 72 hours at a rate of approximately 4.2 L/min and a pressure of 5 psi. The filters collect the planktonic bacterial population which we assume is representative of the entire community (see Metagenomic Analysis). DOC radiocarbon samples were collected after the filter in the outflow near the end of filtering after the majority of water had passed through the filters. Immediately upon completion of filtering the filters were placed in resealable plastic bags on dry ice. The filters remained on dry ice until returning to the lab and then placed at -20°C or -80°C until DNA extraction.

Filter Extraction

DNA was then extracted using a modified QIAGEN Genomic-tip 100/G method. The filters were cut into ~1 cm^2 squares, placed in 50 ml tubes, wetted with TE, rapidly freeze-thawed and immersed in B1 buffer (50 mM Tris-Cl pH 8.0, 50 mM EDTA pH 8.0, 0.5% Tween-20m, 0.5% Triton X-100). Samples were then subject to sonication using a Sonifier Cell Disruptor (Branson Sonic Power Co.) at output control 4 and percent duty cycle 70 followed by repeated vacuum infiltration. Lysozyme, Proteinase-K, and RNase were added to the samples, which were then incubated for at least 1 hr at 37˚C and 200rpm. B2 buffer (3M guanidine HCl, 20% Tween 20) was added, the samples were incubated for at least 1 hr at 50˚C and 200 rpm (alternatively, samples would be incubated for >2hr at 37˚C). Following incubation the

supernatant was thoroughly decanted from the filter paper. DNA was precipitated from the supernatant with 0.7x isopropanol and washed with 70% ethanol.

DNA Purification

DNA was precipitated with 6% PEG8000 and washed with 70% ethanol. DNA was further purified with phenol:chloroform:isoamyl alcohol (25:24:1) and chloroform, followed by ethanol precipitation. The DNA was spun by isopycnic ultracentrifugation in a 1.7 g/ml CsCl solution in a VTi 65.2 Rotor at 45,000 rpm for 48 hr at 25°C. The DNA was removed from the centrifuge tubes in fractions. Each fraction was diluted, precipitated with ethanol, washed in 70% ethanol and suspended in TE (*30*). The DNA purity and concentration were assessed by absorbance and the purest fractions were combined. In preparation for radiocarbon analysis the DNA samples were precipitated with ethanol, washed with 70% ethanol and suspended in deionized water (DI). The samples were then repeatedly lyophilized to dryness and resuspended in DI to remove any traces of ethanol. DNA amounts were determined by absorbance on a Nanodrop (Nanodrop Products, Wilmington, DE).

AMS

Purified DNA samples suspended in water were transferred to quartz combustion tubes and lyophilized to dryness. AMS isotopic standards NIST SRM 4990C Oxalic Acid II and IAEA C-6 sucrose were dissolved in water and underwent the same drying procedure as the DNA. Excess copper oxide (CuO) was added to each dry DNA sample or standard, tubes were evacuated and sealed with a H_2/O_2 torch. Tubes were placed in a furnace set at 900°C for 3.5 hr to combust all carbon to CO_2. The evolved CO_2 was purified, trapped, and reduced to graphite in

the presence of iron catalyst in individual reactors with magnesium perchlorate water traps (*31, 32*). Graphite targets were measured at the Center for AMS at Lawrence Livermore National Laboratory. Due to small sample sizes, CO_2 splits were not possible and $\delta^{13}C$ could not be measured. The $\delta^{13}C$ value of -23±2 ‰, was used for all calculations. Corrections for background contamination introduced during sample preparation were made following the procedures of Brown and Southon (*33*). The measurement error was determined for each sample and ranged between ± 2‰ and 10‰ (1 SD) $\Delta^{14}C$. All ^{14}C data are reported as decay-corrected $\Delta^{14}C$ following the dominant convention of Stuiver and Polach (*34*).

METAGENOMIC ANALYSIS

A second filter sample was collected from the high arsenic F6-25.4 m well but not enough DNA was present for a date. Molecular analysis of the sample was performed to verify that the DNA obtained after the purification steps for radiocarbon analysis was free from lab type bacterial contaminants and to provide a preliminary phylogenetic characterization of the aquifer. Phylogenetic analysis of well water DNA was performed using multiplex tag-encoded pyrosequencing (*35, 36*) and 454 shotgun metagenomic sequencing. To generate amplicons for pyrosequencing, 100 ng of the DNA was amplified using the eubacterial primers 27f and 1492r (*37*). A second PCR was performed on the PCR product using barcoded fusion primers containing a unique 8 bp code (*38*) and targeting the V3 region of the 16S rRNA gene (*39*). Emulsion PCR (emPCR) was performed and the amplicons were sequenced using both the A and B sequencing primers of the Roche 454 titanium chemistry. Following sequencing and data processing to base calls, the combined sequences in a FASTA file were parsed into individual samples using the Ribosomal Database pyrosequencing pipeline (*40*). The well water sample

sequences were uploaded to MG-RAST for phylogenetic identification and analyses (MG-RAST id 4480464.3). This library had 9,195 sequences after quality control (QC) and 1,362 predicted RNA features. A shotgun metagenomic DNA library was prepared from the same DNA sample. The rapid DNA library preparation, emPCR, and sequencing using titanium chemistry were performed following Roche 454 protocols. After sequencing and data processing to base calls, the signal flowgram files (sff) were uploaded to MG-RAST for annotation and analyses (*40*). After QC, this library (MG-RAST id 4461675.3) had 110,675,166 bp with 258,015 predicted ORFs and 38,750 RNA features.

A phylogenetic overview of the F6 well water sample was obtained by 454 sequencing of eubacterial 16S rRNA gene amplicons and total metagenomic DNA followed by annotation in MG-RAST (Figure 1). This annotation and analyses method only considers sequences that can be annotated in existing databases such as RDP and GenBank. Thus it was not used to estimate the microbial diversity of the community, but it can provide an overview of the microbial community. When comparing the 16S rRNA amplicons to the metagenomic sequences the 16S rRNA approach identified more eubacterial phylotypes (203 species) due to the higher number of annotated sequences than the metagenomic approach (92 species, including archaea and eukaryota). However, analysis of the 16S rRNA sequences did not include archaea and may be subject to amplification biases as indicated by the high percentage (65%) of sequences belonging to the family Comamonadaceae. In both the 16S rRNA amplicon and metagenomic libraries the dominant annotated phylum was Protoebacteria which comprised 50% of the RNA sequences in the metagenome library and 91% of the 16S rRNA amplicon sequences. Within this phylum a family found in both the 16S rRNA amplicon and metagenomic libraries was the Rhodocyclaceae family (13% and 17%, respectively) of which *Dechloromonas* sp. were the

dominant species (Figure 1). *Dechlormonas aromatica* is associated with benzene degradation, indicating utilization of complex carbon sources could be an important process and recent studies have also linked benzene degradation to Fe(III) reduction (*41, 42*). Analysis of the metagenomic sequences indicated that autotrophic methanogens, identified as the following species: *Methanobrevibacter smithii* 98% similarity, *Methanothermobacter thermautotrophicus* 96% similarity, *Methanospirillum hungatei* 100% similarity, and *Methanosaeta thermophile* 100% similarity represented 3% of the total RNA sequences identified

The shotgun metagenomic library was analyzed using the SEED subsystem available on MG-RAST server with an E-Value less than $1x10^{-5}$, 50% identity and 50 bp length (Figure 1) (*43*). The most common subsystems were clustering based (14.3%), carbohydrates, (11.3%), amino acids and derivatives (10.0%), and protein metabolism (8.8%). The clustering based subsystem is associated with cell replication and division. The carbohydrate subsystem is associated with carbon metabolism, including Central Carbohydrate metabolism, one-carbon metabolism, and fermentation.

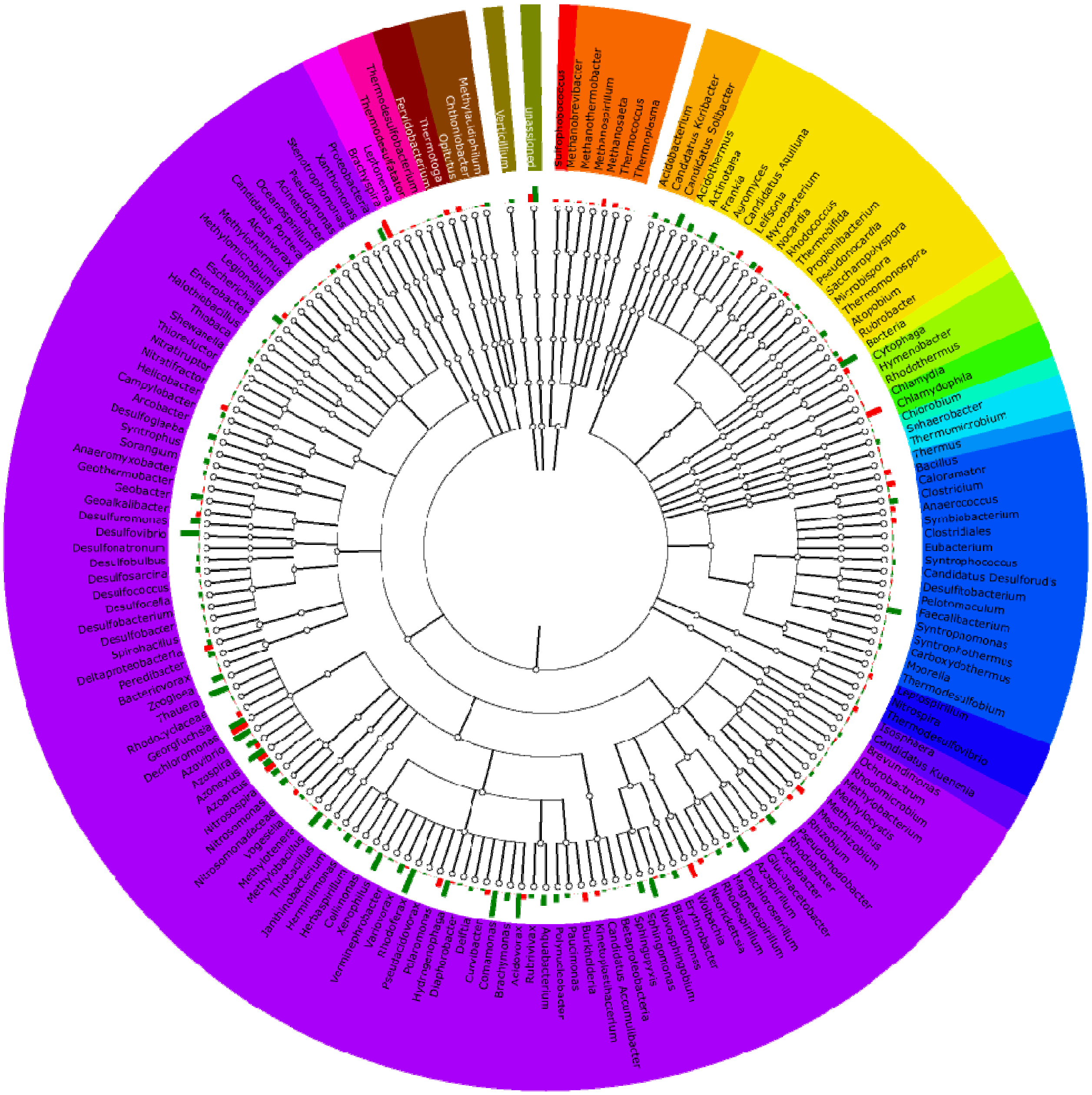

Fig S1: Phylogenetic classification of bacteria to the genus level in well F6 by analyses of 12,605 RNA sequences from metagenomic shotgun sequencing (red) and 9,195 16S rRNA partial gene sequences from tag encoded pyrosequencing (green). The genera are colored by phylum Phylogenetic classification was performed using MG-RAST. The parameters used were based on the Ribosomal Database (RDP release 10) with a maximum e-value of e-10, a minimum identity of 80% similarity, and a minimum alignment length of 50.